10/087,035 EXHIBIT A

Docket No.: 3392

PROVISIONAL PATENT APPLICATION

RAPID FLEXIBLE CONTENT ARRAY AND ONLINE ORDERING SYSTEM

Inventors:

Michael Mittmann, a citizen of the
United States of America
residing at: 2377 St. Francis Drive
Palo Alto, CA 94303

Earl Hubbell, a citizen of the
United States of America
residing at: 416 S. Genesee
Los Angeles, CA 90036

David Smith, a citizen of the
United States of America
residing at: 2034 Finley Place
Santa Clara, CA 95050

Assignee

Affymetrix, Inc.
3380 Central Expressway
Santa Clara, CA 95051

RAPID FLEXIBLE CONTENT ARRAY AND ONLINE ORDERING SYSTEMS

RELATED APPLICATIONS

This application is related to U.S. Patent Application Serial Number 09/721,042, filed on November 21, 2000, entitled "Methods and Computer Software Products for Predicting Nucleic Acid Hybridization Affinity"; U.S. Patent Application Serial Number 09/718,295, filed on November, 21, 2000, entitled "Methods and Computer Software Products for Selecting Nucleic Acid Probes", U.S. Patent Application Serial Number __/____, attorney docket number 3373.1, filed on 12/21, 2000, entitled "Methods For Selecting Nucleic Acid Probes" and U.S. Patent Application Serial Number __/___, attorney docket number 3262.1, filed on January 16, 2001, entitled "Computer Software for Sequence Selection" All the cited applications are incorporated herein by reference in their entireties for all purposes.

FIELD OF INVENTION

This invention is related to DNA micro-array design, fabrication and electronic commerce.

BACKGROUND OF THE INVENTION

U.S. Patent No. 5,424,186 describes a pioneering technique for, among other things, forming and using high density arrays of molecules such as oligonucleotides, RNA or DNA), peptides, polysaccharides, and other materials. This patent is hereby

incorporated by reference for all purposes. However, there is still great need for methods, systems and software for designing, fabricating and commercializing high density nucleic acid probe arrays.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and form a part of this specification, illustrate embodiments of the invention and, together with the description, serve to explain the principles of the invention:

FIGURE 1 illustrates one embodiment of the offset strategy of the invention.

FIGURE 2 shows information density of possible offsets strategies.

FIGURE 3 compares offset vs shift mask strategy.

FIGURE 4 shows one embodiment of the flexible content array offering process of the invention.

FIGURE 5 shows one embodiment of the flexible content array offering process of the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Reference will now be made in detail to the preferred embodiments of the invention. While the invention will be described in conjunction with the preferred embodiments, it will be understood that they are not intended to limit the invention to these embodiments. On the contrary, the invention is intended to cover alternatives, modifications and equivalents, which may be included within the spirit and scope of the invention. All cited references, including patent and non-patent literature, are incorporated herein by reference in their entireties for all purposes.

I. High Density Probe Arrays

The methods, computer software and systems of the invention are particularly useful for designing high density nucleic acid probe arrays.

High density nucleic acid probe arrays, also referred to as "DNA Microarrays," have become a method of choice for monitoring the expression of a large number of

genes and for detecting sequence variations, mutations and polymorphism. As used herein, "nucleic acids" may include any polymer or oligomer of nucleosides or nucleotides (polynucleotides or oligonucleotidies), which include pyrimidine and purine bases, preferably cytosine, thymine, and uracil, and adenine and guanine, respectively. See Albert L. Lehninger, PRINCIPLES OF BIOCHEMISTRY, at 793-800 (Worth Pub. 1982) and L. Stryer, BIOCHEMISTRY, 4th Ed. (March 1995), both incorporated by reference. "Nucleic acids" may include any deoxyribonucleotide, ribonucleotide or peptide nucleic acid component, and any chemical variants thereof, such as methylated, hydroxymethylated or glucosylated forms of these bases, and the like. The polymers or oligomers may be heterogeneous or homogeneous in composition, and may be isolated from naturally-occurring sources or may be artificially or synthetically produced. In addition, the nucleic acids may be DNA or RNA, or a mixture thereof, and may exist permanently or transitionally in single-stranded or double-stranded form, including homoduplex, heteroduplex, and hybrid states.

"A target molecule" refers to a biological molecule of interest. The biological molecule of interest can be a ligand, receptor, peptide, nucleic acid (oligonucleotide or polynucleotide of RNA or DNA), or any other of the biological molecules listed in U.S. Pat. No. 5,445,934 at col. 5, line 66 to col. 7, line 51, which is incorporated herein by reference for all purposes. For example, if transcripts of genes are the interest of an experiment, the target molecules would be the transcripts. Other examples include protein fragments, small molecules, etc. "Target nucleic acid" refers to a nucleic acid (often derived from a biological sample) of interest. Frequently, a target molecule is detected using one or more probes. As used herein, a "probe" is a molecule for detecting a target molecule. It can be any of the molecules in the same classes as the target referred to above. A probe may refer to a nucleic acid, such as an oligonucleotide, capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (i.e. A, G, U, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, the bases in probes may be

joined by a linkage other than a phosphodiester bond, so long as the bond does not interfere with hybridization. Thus, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. Other examples of probes include antibodies used to detect peptides or other molecules, any ligands for detecting its binding partners. When referring to targets or probes as nucleic acids, it should be understood that these are illustrative embodiments that are not to limit the invention in any way.

In preferred embodiments, probes may be immobilized on substrates to create an array. An "array" may comprise a solid support with peptide or nucleic acid or other molecular probes attached to the support. Arrays typically comprise a plurality of different nucleic acids or peptide probes that are coupled to a surface of a substrate in different, known locations. These arrays, also described as "microarrays" or colloquially "chips" have been generally described in the art, for example, in Fodor et al., *Science*, 251:767-777 (1991), which is incorporated by reference for all purposes. Methods of forming high density arrays of oligonucleotides, peptides and other polymer sequences with a minimal number of synthetic steps are disclosed in, for example, U.S. Pat. Nos. 5,143,854, 5,252,743, 5,384,261, 5,405,783, 5,424,186, 5,429,807, 5,445,943, 5,510,270, 5,677,195, 5,571,639, 6,040,138, all incorporated herein by reference for all purposes. The oligonucleotide analogue array can be synthesized on a solid substrate by a variety of methods, including, but not limited to, light-directed chemical coupling, and mechanically directed coupling. See Pirrung et al., U.S. Pat. No. 5,143,854 (see also PCT Application No. WO 90/15070) and Fodor et al., PCT Publication Nos. WO 92/10092 and WO 93/09668, U.S. Pat. Nos. 5,677,195, 5,800,992 and 6,156,501, which disclose methods of forming vast arrays of peptides, oligonucleotides and other molecules using, for example, light-directed synthesis techniques. See also, Fodor, et al., *Science*, 251, 767-77 (1991). These procedures for synthesis of polymer arrays are now referred to as VLSIPS™ procedures.

Methods for making and using molecular probe arrays, particularly nucleic acid probe arrays are also disclosed in, for example, U.S. Pat. Nos. 5,143,854, 5,242,974,

5,252,743, 5,324,633, 5,384,261, 5,405,783, 5,409,810, 5,412,087, 5,424,186, 5,429,807, 5,445,934, 5,451,683, 5,482,867, 5,489,678, 5,491,074, 5,510,270, 5,527,681, 5,527,681, 5,541,061, 5,550,215, 5,554,501, 5,556,752, 5,556,961, 5,571,639, 5,583,211, 5,593,839, 5,599,695, 5,607,832, 5,624,711, 5,677,195, 5,744,101, 5,744,305, 5,753,788, 5,770,456, 5,770,722, 5,831,070, 5,856,101, 5,885,837, 5,889,165, 5,919,523, 5,922,591, 5,925,517, 5,658,734, 6,022,963, 6,150,147, 6,147,205, 6,153,743 and 6,140,044, all of which are incorporated by reference in their entireties for all purposes.

Microarray can be used in a variety of ways. A preferred microarray contains nucleic acids and is used to analyze nucleic acid samples. Typically, a nucleic acid sample is prepared from appropriate source and labeled with a signal moiety, such as a fluorescent label. The sample is hybridized with the array under appropriate conditions. The arrays are washed or otherwise processed to remove non-hybridized sample nucleic acids. The hybridization is then evaluated by detecting the distribution of the label on the chip. The distribution of label may be detected by scanning the arrays to determine fluorescence intensity distribution. Typically, the hybridization of each probe is reflected by several pixel intensities. The raw intensity data may be stored in a gray scale pixel intensity file. The GATC™ Consortium has specified several file formats for storing array intensity data. The final software specification is available at www.gatcconsortium.org and is incorporated herein by reference in its entirety. The pixel intensity files are usually large. For example, a GATC™ compatible image file may be approximately 50 Mb if there are about 5000 pixels on each of the horizontal and vertical axes and if a two byte integer is used for every pixel intensity. The pixels may be grouped into cells (see, GATC™ software specification). The probes in a cell are designed to have the same sequence (i.e., each cell is a probe area). A CEL file contains the statistics of a cell, e.g., the 75th percentile and standard deviation of intensities of pixels in a cell. The 50, 60, 70, 75 or 80th percentile of pixel intensity of a cell is often used as the intensity of the cell.

Nucleic acid probe arrays have found wide applications in gene expression monitoring, genotyping and mutation detection. For example, massive parallel gene

expression monitoring methods using nucleic acid array technology have been developed to monitor the expression of a large number of genes (e.g., U.S. Patent Numbers 5,871,928, 5,800,992 and 6,040,138; de Saizieu et al., 1998, Bacteria Transcript Imaging by Hybridization of total RNA to Oligonucleotide Arrays, NATURE BIOTECHNOLOGY, 16:45-48; Wodicka et al., 1997, Genome-wide Expression Monitoring in Saccharomyces cerevisiae, NATURE BIOTECHNOLOGY 15:1359-1367; Lockhart et al., 1996, Expression Monitoring by Hybridization to High Density Oligonucleotide Arrays. NATURE BIOTECHNOLOGY 14:1675-1680; Lander, 1999, Array of Hope, NATURE-GENETICS, 21(suppl.), at 3, all incorporated herein by reference for all purposes). Hybridization-based methodologies for high throughput mutational analysis using high-density oligonucleotide arrays (DNA chips) have been developed, see Hacia et al., 1996, Detection of heterozygous mutations in BRCA1 using high density oligonucleotide arrays and two-color fluorescence analysis. Nat. Genet. 14:441-447, Hacia et al., New approaches to BRCA1 mutation detection, Breast Disease 10:45-59 and Ramsey 1998, DNA chips: State-of-Art, Nat Biotechnol. 16:40-44, all incorporated herein by reference for all purposes). Oligonucleotide arrays have been used to screen for sequence variations in, for example, the CFTR gene (U.S. Patent Number 6,027,880, Cronin et al., 1996, Cystic fibrosis mutation detection by hybridization to light-generated DNA probe arrays. Hum. Mut. 7:244-255, both incorporated by reference in their entireties), the human immunodeficiency virus (HIV-1) reverse transcriptase and protease genes (U.S. Patent Number 5,862,242 and Kozal et al., 1996, Extensive polymorphisms observed in HIV-1 clade B protease gene using high density oligonucleotide arrays. Nature Med. 1:735-759, both incorporated herein by reference for all purposes), the mitochondrial genome (Chee et al., 1996, Accessing genetic information with high density DNA arrays. Science 274:610-614) and the BRCA1 gene (U.S. Patent Number 6,013,449, incorporated herein by reference for all purposes).

Methods for signal detection and processing of intensity data are additionally disclosed in, for example, U.S. Pat. Nos. 5,445,934, 547,839, 5,578,832, 5,631,734, 5,800,992, 5,856,092, 5,936,324, 5,981,956, 6,025,601, 6,090,555, 6,141,096, 6,141,096,

and 5,902,723. Methods for array based assays, computer software for data analysis and applications are additionally disclosed in, e.g., U.S. Pat. Nos. 5,527,670, 5,527,676, 5,545,531, 5,622,829, 5,631,128, 5,639,423, 5,646,039, 5,650,268, 5,654,155, 5,674,742, 5,710,000, 5,733,729, 5,795,716, 5,814,450, 5,821,328, 5,824,477, 5,834,252, 5,834,758, 5,837,832, 5,843,655, 5,856,086, 5,856,104, 5,856,174, 5,858,659, 5,861,242, 5,869,244, 5,871,928, 5,874,219, 5,902,723, 5,925,525, 5,928,905, 5,935,793, 5,945,334, 5,959,098, 5,968,730, 5,968,740, 5,974,164, 5,981,174, 5,981,185, 5,985,651, 6,013,440, 6,013,449, 6,020,135, 6,027,880, 6,027,894, 6,033,850, 6,033,860, 6,037,124, 6,040,138, 6,040,193, 6,043,080, 6,045,996, 6,050,719, 6,066,454, 6,083,697, 6,114,116, 6,114,122, 6,121,048, 6,124,102, 6,130,046, 6,132,580, 6,132,996 and 6,136,269, all of which are incorporated by reference in their entireties for all purposes.

Nucleic acid probe array technology, use of such arrays, analysis array based experiments, associated computer software, composition for making the array and practical applications of the nucleic acid arrays are also disclosed, for example, in the following U.S. Patent Applications: 07/838,607, 07/883,327, 07/978,940, 08/030,138, 08/082,937, 08/143,312, 08/327,522, 08/376,963, 08/440,742, 08/533,582, 08/643,822, 08/772,376, 09/013,596, 09/016,564, 09/019,882, 09/020,743, 09/030,028, 09/045,547, 09/060,922, 09/063,311, 09/076,575, 09/079,324, 09/086,285, 09/093,947, 09/097,675, 09/102,167, 09/102,986, 09/122,167, 09/122,169, 09/122,216, 09/122,304, 09/122,434, 09/126,645, 09/127,115, 09/132,368, 09/134,758, 09/138,958, 09/146,969, 09/148,210, 09/148,813, 09/170,847, 09/172,190, 09/174,364, 09/199,655, 09/203,677, 09/256,301, 09/285,658, 09/294,293, 09/318,775, 09/326,137, 09/326,374, 09/341,302, 09/354,935, 09/358,664, 09/373,984, 09/377,907, 09/383,986, 09/394,230, 09/396,196, 09/418,044, 09/418,946, 09/420,805, 09/428,350, 09/431,964, 09/445,734, 09/464,350, 09/475,209, 09/502,048, 09/510,643, 09/513,300, 09/516,388, 09/528,414, 09/535,142, 09/544,627, 09/620,780, 09/640,962, 09/641,081, 09/670,510, 09/685,011, and 09/693,204 and in the following Patent Cooperative Treaty (PCT) applications/publications: PCT/NL90/00081, PCT/GB91/00066, PCT/US91/08693, PCT/US91/09226, PCT/US91/09217, WO/93/10161, PCT/US92/10183, PCT/GB93/00147, PCT/US93/01152, WO/93/22680,

PCT/US93/04145, PCT/US93/08015, PCT/US94/07106, PCT/US94/12305, PCT/GB95/00542, PCT/US95/07377, PCT/US95/02024, PCT/US96/05480, PCT/US96/11147, PCT/US96/14839, PCT/US96/15606, PCT/US97/01603, PCT/US97/02102, PCT/GB97/005566, PCT/US97/06535, PCT/GB97/01148, PCT/GB97/01258, PCT/US97/08319, PCT/US97/08446, PCT/US97/10365, PCT/US97/17002, PCT/US97/16738, PCT/US97/19665, PCT/US97/20313, PCT/US97/21209, PCT/US97/21782, PCT/US97/23360, PCT/US98/06414, PCT/US98/01206, PCT/GB98/00975, PCT/US98/04280, PCT/US98/04571, PCT/US98/05438, PCT/US98/05451, PCT/US98/12442, PCT/US98/12779, PCT/US98/12930, PCT/US98/13949, PCT/US98/15151, PCT/US98/15469, PCT/US98/15458, PCT/US98/15456, PCT/US98/16971, PCT/US98/16686, PCT/US99/19069, PCT/US98/18873, PCT/US98/18541, PCT/US98/19325, PCT/US98/22966, PCT/US98/26925, PCT/US98/27405 and PCT/IB99/00048, all the above cited patent applications and other references cited throughout this specification are incorporated herein by reference in their entireties for all purposes.

II. Mask Offset Strategy

In one aspect of the invention, an offset strategy is provided. The strategy better utilizes the area of each mask for photodirected synthesis of oligonucleotides. This offset strategy allows each mask to contain several reticles per layer thus reducing the cost of the mask set.

In some preferred embodiments, the offset strategy allows an N x M array of reticles to be positioned on each mask layer. By using the area of that array plus the bordered blank areas between arrays, the offset strategy allows for a larger print area. Figure 1 illustrates a 2 x 2 reticle mask strategy using the space between arrays and the active area of the array. In this case, a 75-cycle synthesis can be performed with a 19 layer mask set – approximately a 4-fold reduction of mask cost. Within this strategy the wafer moves but the mask does not. To fully utilize this strategy with existing Genechip® array (Affymetrix, Santa Clara, CA) manufacturing equipment, 4 alignment marks may be required on the substrate – one for each of the reticles. Likewise, if other

offset strategies are used, a new substrate would be required that contained the appropriate number of marks. Table 1 represents several strategies utilizing the offset strategies and Table 2 shows the active areas of these designs. Multiple alignment marks allow for a more efficient synthesis process as shown below.

Offset Photolithography Process

Offset (~3000 – 4000 m) → Align (1 alignment) → Expose (h)

Shift Mask Photolithography Process

Align (3 alignments) → Shift (~30– 40 m) → Expose (h)

The Offset strategy has several benefits over the current Shift mask strategy:

1. Though the offset strategy may require a ~ 100-fold larger wafer stage move, by aligning after the move, systematic equipment errors will be minimized and alignment quality will be consistent with current commercial products. Shift mask arrays require a "blind" move that adds systematic equipment error to existing manual alignment tolerances. Recent results verify that current manufacturing equipment is incapable of large moves required in the offset strategy.

2. To manufacture probe arrays containing 25mers with shift masks (~9 masks for PM-MM strategy), 105 synthesis cycles and 315 mask alignments are required. To manufacture probe arrays containing 25mers with the offset strategy (~5-19 masks required), 75 cycles and 75 alignments are required. Alignment after the move will ensure quality, while fewer alignments will provide better throughput.

3. Though offset arrays will have an active area smaller than the dimensions inside the current picture frame on the chip, the active area will have corner checkerboards and will allow for the manufacture of current probe sets from existing arrays to be built on a sub-set array. Offset arrays will have a similar look

and feel to our commercial arrays. Shift mask arrays cannot contain subsets of current probe sets and would have a different look and feel including a requirement of manual corner alignment.

4. Offset masks will allow ~ 1000 probe sets to be represented within a midi format (100) array. Shift strategies will allow ~800 probe sets to be represented. The shift strategy would tile ~200 contiguous bases in the standard PM-MM expression layout.

To implement the superior offset strategy, 3 manufacturing efforts may be needed:

1. Hardware modifications – a larger window is required with the wafer holder (chuck) on the Odo photolithography station to allow alignment of the multiple alignment marks on the substrate. To visualize this larger area, a more diffuse backlight is required. Sigma photolithography stations (Operational late Q2) were designed to with this requirement.

2. Software Modifications – Manufacturing software is required to allow for multiple alignment marks within a substrate and to increase operational efficiencies.

3. New substrates – For each new format, i.e. 2 X 2 for micro arrays(400) and 3 X 3 for midi arrays (100), a new substrate with appropriate alignment marks are required. Each new substrate will require a separate part number and an inventory to allow rapid turnaround.

Table 1 – Information Density of Possible Offsets Strategies (16 pp per probe set)

Note: Excludes control probes and control probe sets (~2000 probes, 120-150 probe sets)

x division s	y divisions	Total Genes for Std Format (49)	Total Genes for Midi Format (100)	Total Genes for MiniFormat (169)	Total Genes for MicroFormat (400)	number masks needed
2	2	4875	2380	1326	496	19
2	3	3246	1587	881	362	13
2	4	2431	1190	663	271	10
3	3	2161	1058	586	264	9
3	5	1298	632	351	158	5
4	5	972	474	263	118	4
5	5	780	378	210	94	3
5	8	483	237	130	58	2
5	15	256	123	69	30	1

Table 2 – Active Area Dimensions for Possible Offsets Strategies (m)

		Std Format (49)	Midi Format (100)	Mini Format (169)	Micro Format (400)	number masks needed
Synthesis Area		12800 m	8060 m	5250 m	2530 m	
Space Between Picture Frames		3000 m	3000 m	3000 m	3000 m	
x	2	7900	5530	4125	2765*	19
y	2	7900	5530	4125	2765*	
x	2	7900	5530	4125	2765*	13
y	3	5267	3687	2750	1843	
x	2	7900	5530	4125	2765*	10
y	4	3950	2765	2063	1383	
x	3	5267	3687	2750	1843	9
y	3	5267	3687	2750	1843	
x	3	5267	3687	2750	1843	5
y	5	3160	2212	1650	1106	
x	4	3950	2765	2063	1383	4
y	5	3160	2212	1650	1106	

x	5	3160	2212	1650	1106	3
y	5	3160	2212	1650	1106	
x	5	3160	2212	1650	1106	2
y	8	1975	1383	1031	691	
x	5	3160	2212	1650	1106	1
y	15	1053	737	550	369	

* - Current maximum area for micro array is 2530 µm

III. Order-Design Process

Figures 4 and 5 outline embodiments of the Order – Design Process of the invention. The process may be implemented in conjunction with and as a part of an electronic commerce portal for genetic researchers. Genomic web portals are disclosed in, e.g., U.S. Patent Application Serial Number __/____, Attorney Docket Number 3291.1, filed on January 24, 2001, entitled "METHOD, SYSTEM AND COMPUTER SOFTWARE FOR PROVIDING A GENOMIC WEB PORTAL", which is incorporated herein by reference for all purposes.

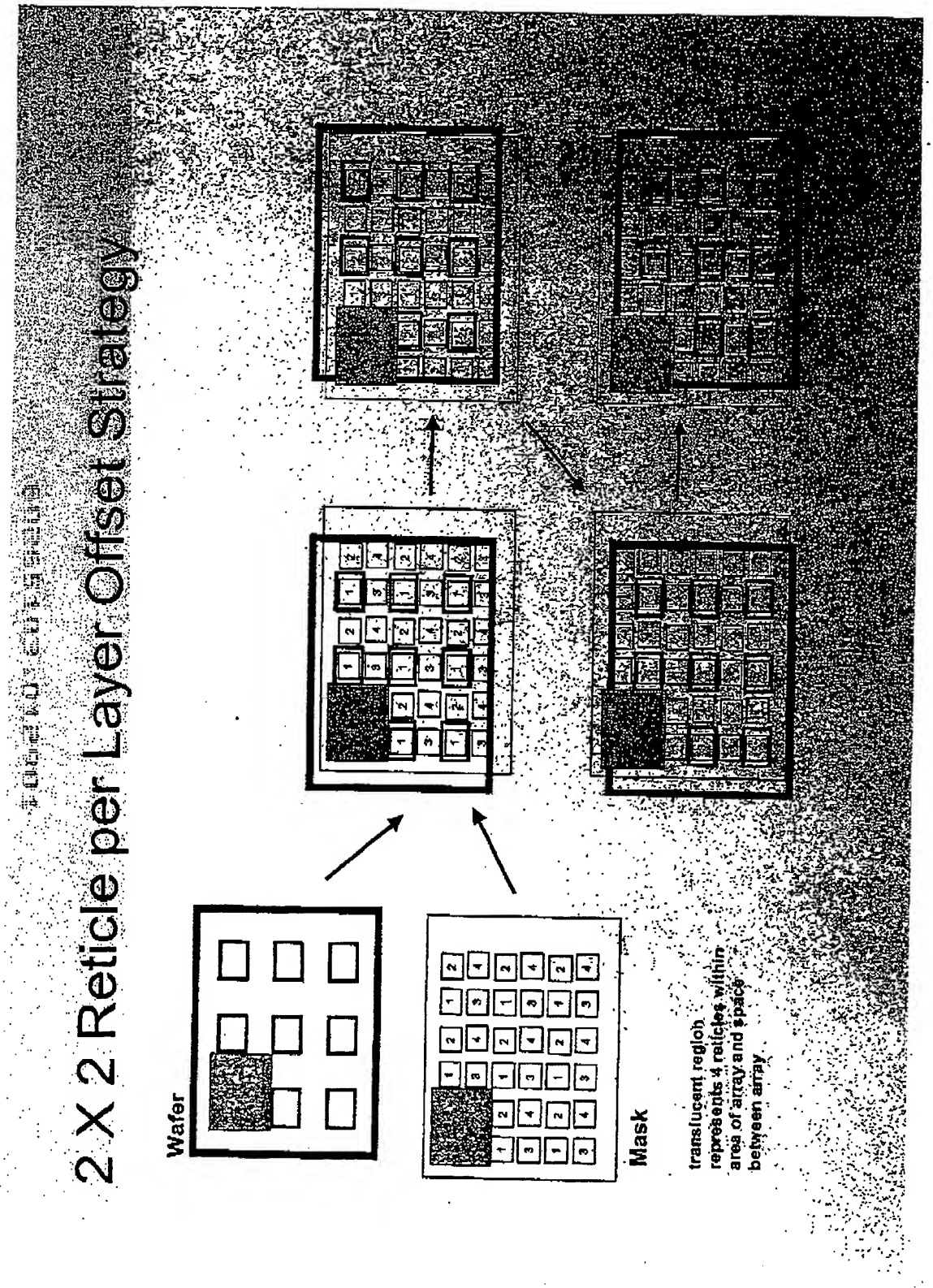
In other preferre embodiments, the process may involve an electronic form submitted through e-mail with special processes in place for the processing / design applications to convert this file to the customer's design. This "checker" program may be used to verify an acceptable order and confirm a customer's commitment to that order. FIGURE 5 shows such a process which include "checker" , chip design templates, subset application, and a library file generation application.

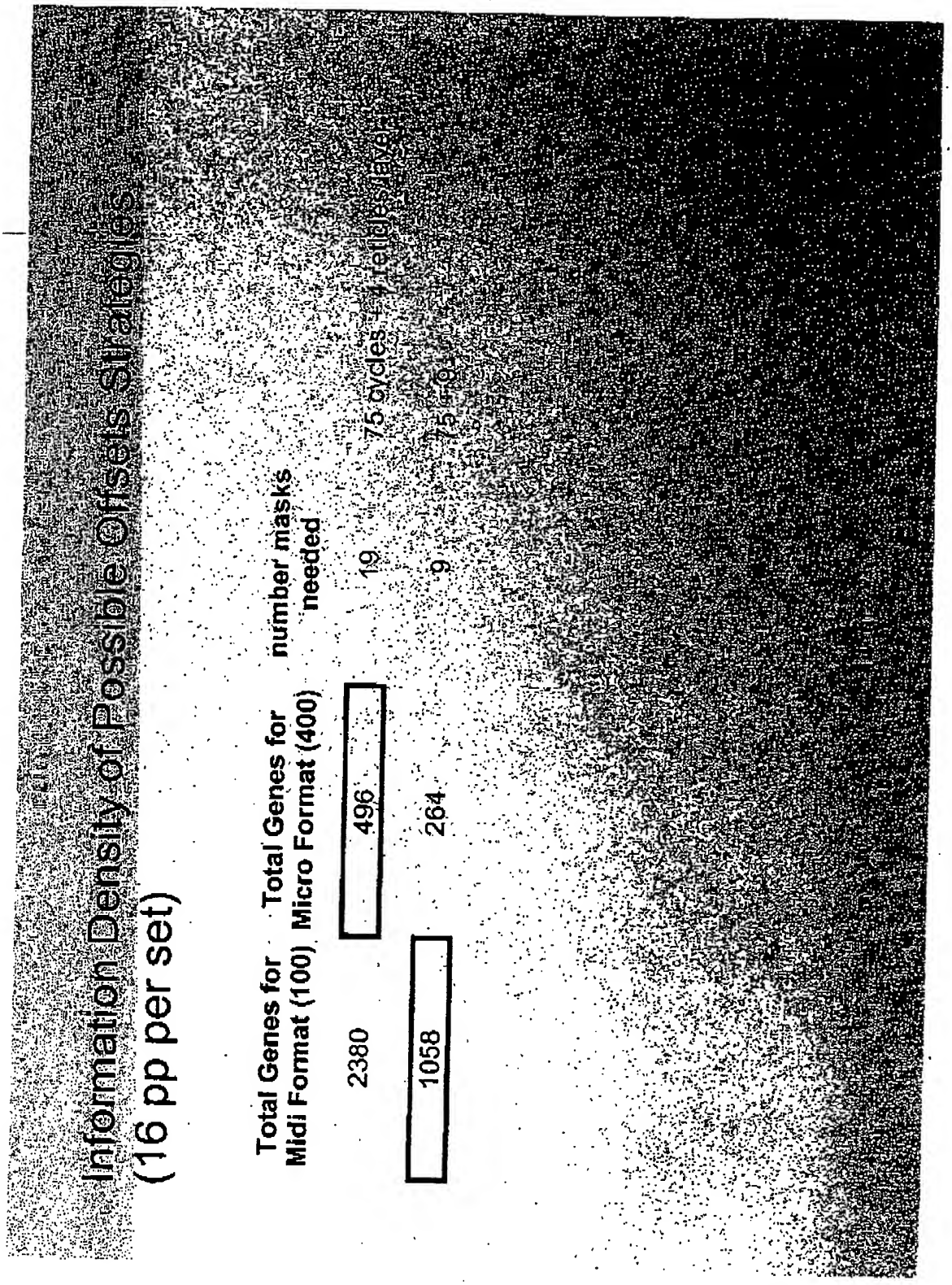
In preferred embodiments, the order processing, design, mask vendor management, array manufacturing, and shipment are integrated.

Methods, software and systems for conducting electronic commerce and for automatic processing of orders are well known to those with ordinary skill in the art.

The present invention provides methods, systems and computer software products for nucleic acid probe array design. It is to be understood that the above description is intended to be illustrative and not restrictive. Many variations of the invention will be apparent to those of skill in the art upon reviewing the above description. All cited

references, including patent and non-patent literature, are incorporated herein by reference in their entireties for all purposes.


2 X 2 Reticle per Layer Offset Strategy
Wafer
Mask
translucent region represents 4 reticles within area of array and space between array


Information Density of Possible Offsets Strategies
(16 pp per set)
Total Genes for Midi Format (100)
Total Genes for Micro Format (400)
number masks needed
2380
1058
496
264
19
9
75 cycles / reticle/ave
75 cycles / reticle/ave

Offset vs Shift Mask Strategy

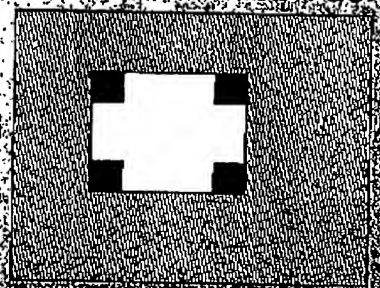
Offset Photolithography Process

Offset (~3000 – 4000 μm) → Align (1 alignment) → Expose (hv)

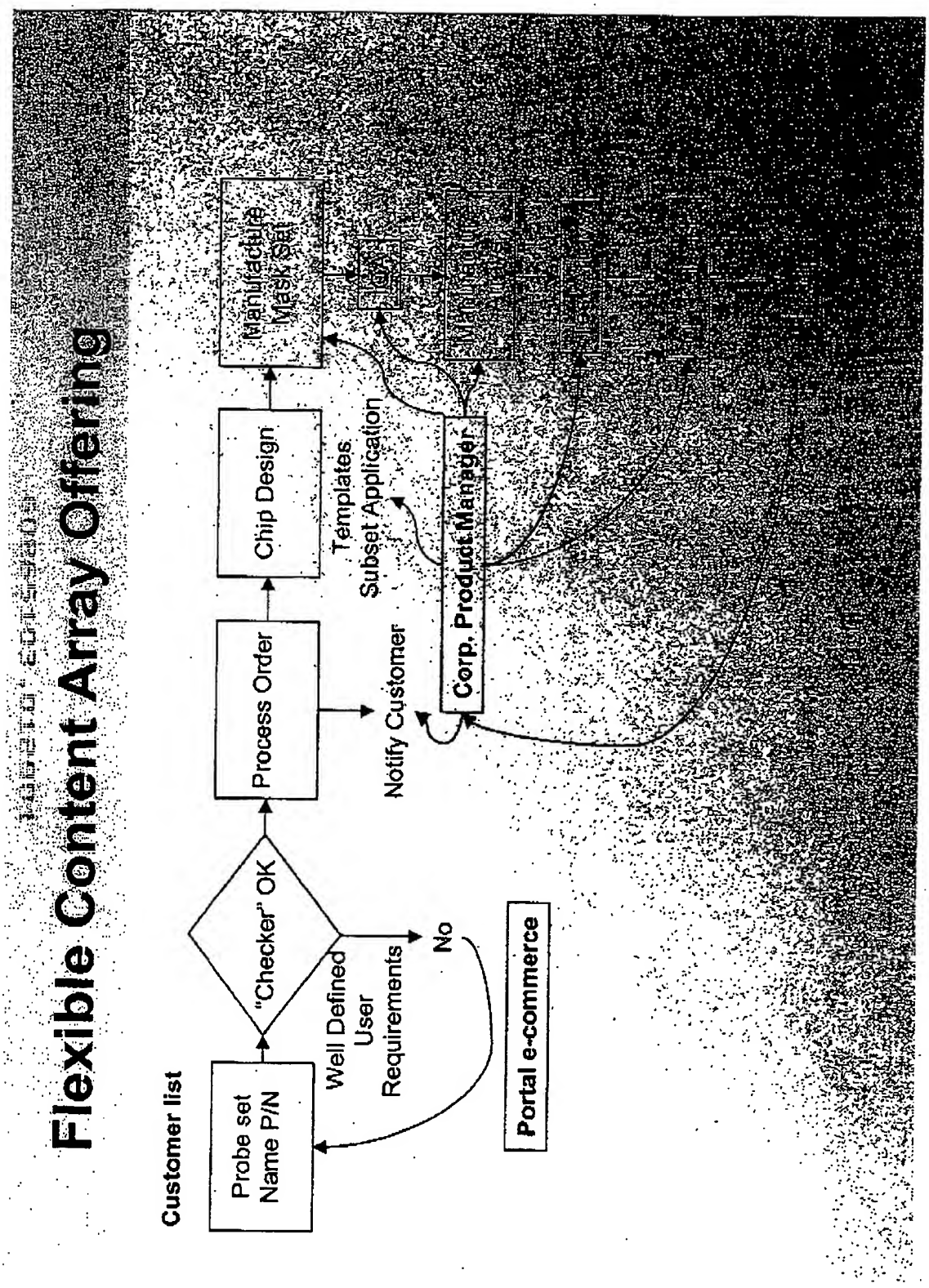
Shift Mask Photolithography Process

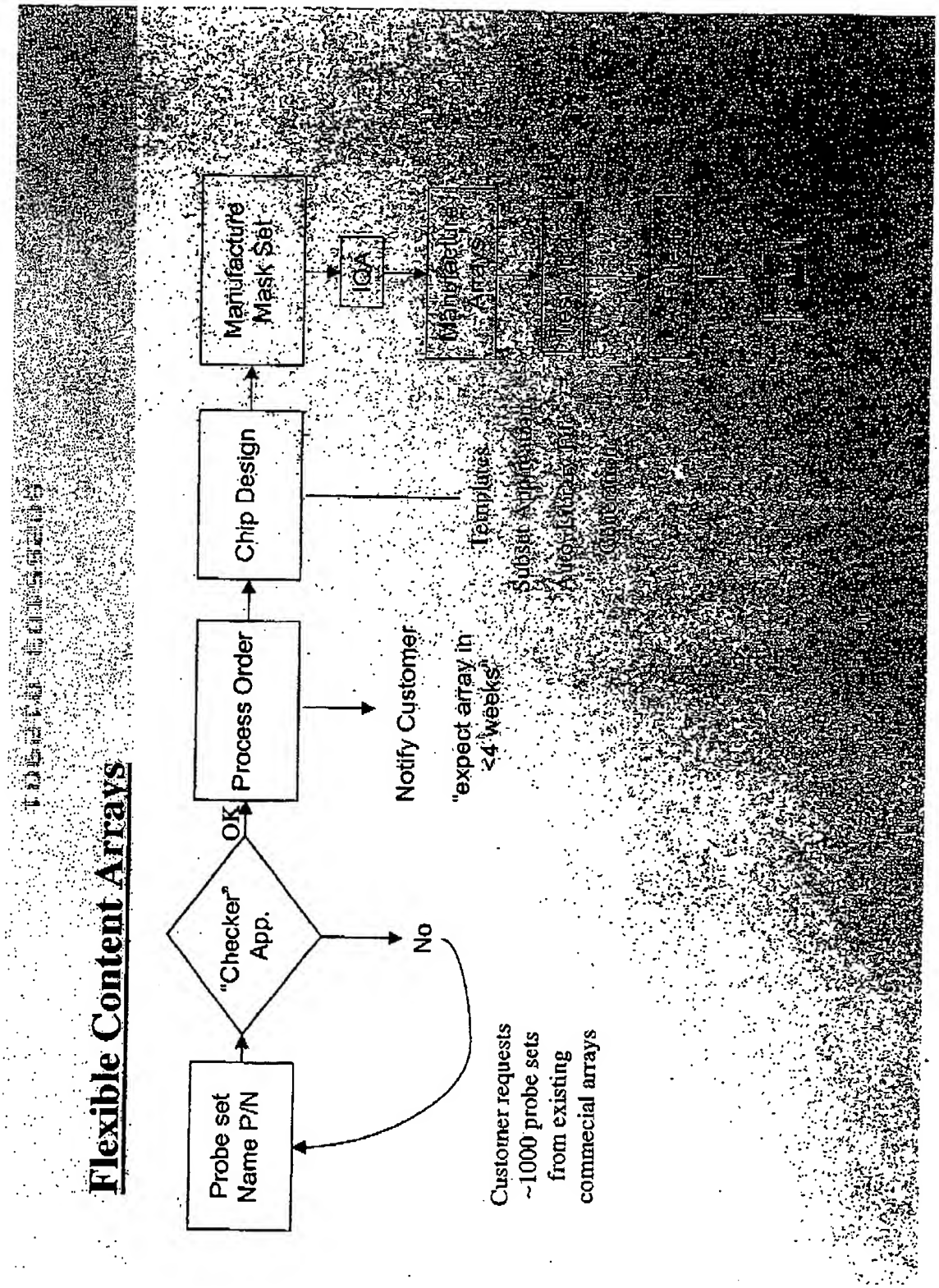
Align (3 alignments) → Shift (~30– 40 μm) → Expose (hv)

1. ~100-fold larger moves require post-move alignment
2. Offset - 75 synthesis cycles vs Shift - 105
3. Offset Strategy would have similar look as custom array (complementary checkerboards)



4. Offset - subset arrays vs Shift - different [illegible]


Flexible Content Array Offering
Customer list
Probe set Name P/N
"Checker" OK
Well Defined User Requirements
No
Portal e-commerce
Process Order
Chip Design
Manufacture Mask Set
Templates
Subset Application
Notify Customer
Corp. Product Manager


Flexible Content Arrays
Probe set Name P/N
"Checker" App.
OK
No
Process Order
Chip Design
Manufacture Mask Set
IQA
Manufacture Arrays
Notify Customer
"expect array in <4 weeks"
Customer requests ~1000 probe sets from existing commecial arrays